CS 2300 Program 4
Due 22 Nov 19

Administrative:

1. Put your information (name, class, etc) in a header comment block.
2. Please comment appropriately. We are not looking for "production quality" comments but the grader should be able to follow your code. This is particularly important if you code doesn't work properly.

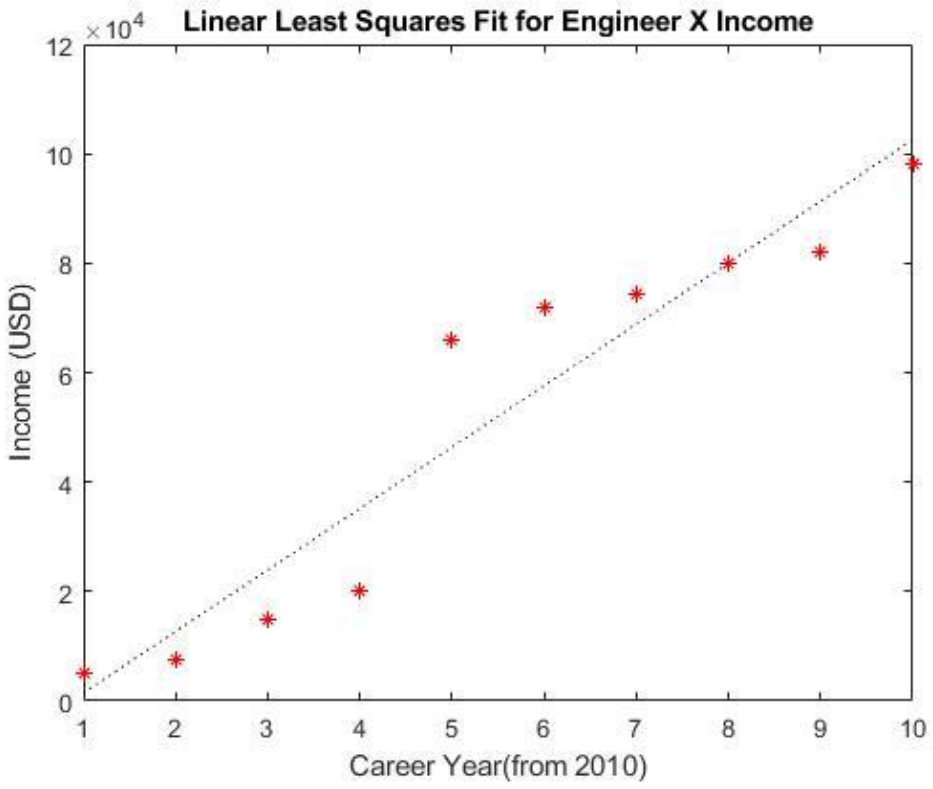
Description of the Assignment:

For this assignment, we are turning the corner on writing low-level matrix and vector functions. For this and the next assignment, you can use built-in functions like matrix multiply, inversion, etc.

As we discussed in class, we can use linear algebra to find a *least-squares fit* for data. For example, assume an engineer's income for his/her college and post-graduation was the following:

Year	Income ($)	Notes	Year in the Workforce
2010	5000	Freshman Year	**1**
2011	7500		**2**
2012	15000		**3**
2013	20000		**4**
2014	66000	Graduation!	**5**
2015	72000		**6**
2016	74500		**7**
2017	80000		**8**
2018	82000		**9**
2019	98000	Promotion!	**10**

What will this engineer make in 2024 (year 15)? Of course, there is no way to know but we can use regression to form an estimate. The numerical value of the year can be used – the math works out - but I've used years in the workforce instead. Here is the linear fit (line) that minimizes the distance between the line and the data points:



The coefficients for this line are [1.123030303030303e+04;-9.766666666666660e+03], meaning the line that is fit is (about):

$$Income(year) = 11230 * year - 9766$$

The coefficients for the linear best fit are found by solving the following system:

$$Ax = b$$

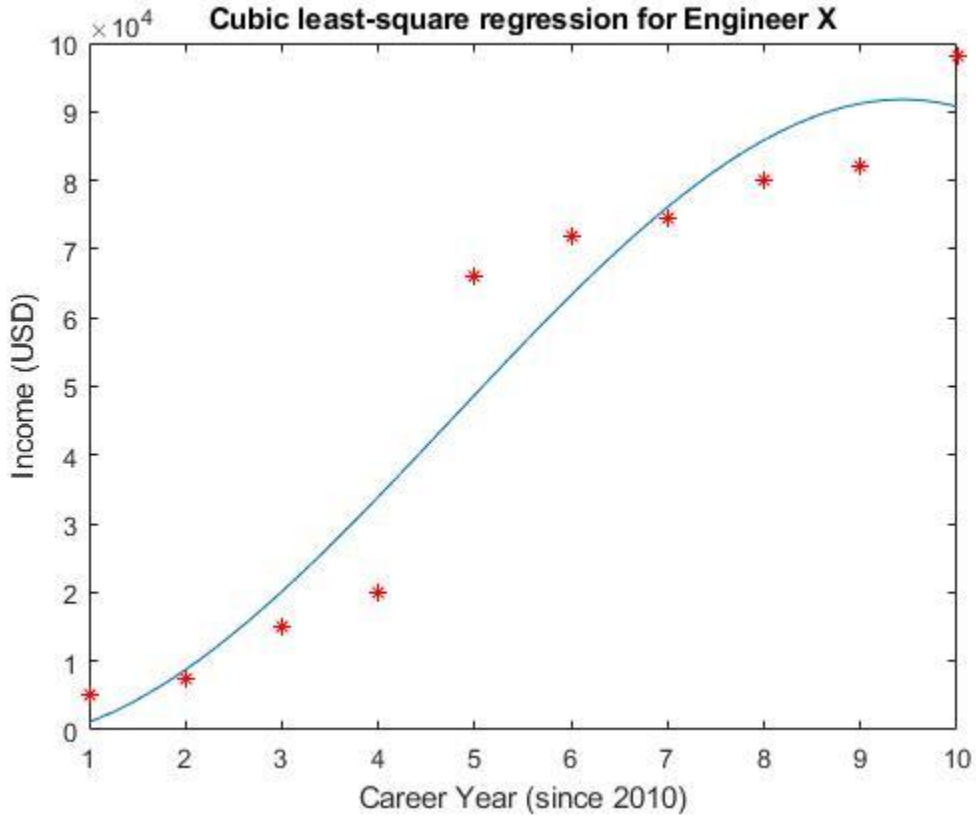
where A, x and b are:

$$\begin{bmatrix} 1 & 1 \\ 2 & 1 \\ 3 & 1 \\ 4 & 1 \\ 5 & 1 \\ 6 & 1 \\ 7 & 1 \\ 8 & 1 \\ 9 & 1 \\ 10 & 1 \end{bmatrix} \begin{bmatrix} x1 \\ x0 \end{bmatrix} = \begin{bmatrix} 5000 \\ 7500 \\ 15000 \\ 20000 \\ 66000 \\ 72000 \\ 74500 \\ 80000 \\ 82000 \\ 98000 \end{bmatrix}$$

First, lets look at the matrix A. The first column is our x values raised to the first power. The second column is our set x raised to the zeroth power. The b vector are the coefficients of the

fit . Finally, the vector b is our y values (or the income). How can we solve this?! The matrices are not square, so finding the inverse is not possible. See section 12.7 of the text for pointer on how to solve this *overdetermined* system.

It makes sense that linear algebra can be used to find a linear "best fit" but it may be less clear that we can find quadratic, cubic, quartic, etc fit using linear algebra as well. Here is an example of a cubic fit for our dataset:



Deliverables:

For this assignment:

1. Implement use the data set above to implement and plot a linear least- squares fit
2. Implement and plot a quadratic least-squares fit
3. For each, use your results to estimate the career income at year 15 years. Does the estimate make sense? Explain why it does or does not.